AD-A049 111 RAND CORP SANTA MONICA CALIF F/G 17/8

REMOTE ACQUISITION OF GROUND TARGETS THROUGH IMAGE CORRELATION --ETC(U)

MAR 77 H H BAILEY

UNCLASSIFIED RAND/P-58 NL

| OF |

AD
A049111

END
DATE
FILMED
3-78
DDC

2

6

REMOTE ACQUISITION OF GROUND TARGETS THROUGH IMAGE CORRELATION (MAP-MATCHING)

10 H. H. Bailey

11 Mar[illegible] [illegible]77

12 34p.

14 RAND/P-58

P-5814

296 600 B

ABSTRACT

A class of weapon guidance based on a technique known as map-matching or image correlation has potential importance in the near future for both strategic and tactical applications in which accurate weapon delivery from standoff is required. The paper presents a brief tutorial explanation of the nature of this guidance technique, discusses both its strengths and its weaknesses, and describes some current research on the acquisition phase (the avoidance of gross errors) in image correlation.

This paper was presented at the Airpower Symposium, "The Impact of Technology on Air Warfare," Air War College (AU), Maxwell Air Force Base, Alabama on 30 March 1977.

I. INTRODUCTION

In any military conflict, the side with numerically inferior forces must pay particular attention to maintaining high effectiveness in its weapons. This means, among other things, achieving sufficiently high accuracy against specified targets that few weapons are wasted. (Don't fire until you see the whites of their eyes.) If, in addition, air supremacy is not available over enemy territory--for example, during the early phases of a conflict--then air-launched weapons may have to be delivered from standoff. This term means that such weapons are "indirectly" fired (to use the Army expression) from a point beyond visual line of sight to the target. But they must still be capable of target-oriented terminal guidance to get the desired accuracy. Thus, there is a strong requirement for achieving autonomous target acquisition and tracking by standoff weapons.

Automatic tracking following acquisition or "lock-on" is a well-developed technology, and the analysis and prediction of tracking performance is also fairly well advanced. Not as much can be said for the acquisition function, however. To date, human intervention, in one form or another, must be involved to identify and acquire ground targets in a cluttered terrain background. There are, nevertheless, two generic approaches to realizing the required self-contained target acquisition on air-delivered weapons.

One approach is a straightforward extrapolation from conventional visual acquisition by the pilot, namely, to place a remote imaging

sensor (TV, IR or even radar) on the weapon, to relay the image so obtained back to the pilot over a radio link, to allow him to find the target on the image displayed to him and then direct the weapon (over a second radio link) to lock onto the appropriate portion of the image and steer itself to impact with the target. Thus, the weapon becomes a form of remotely piloted vehicle (RPV). It utilizes technologies that are known to exist; but the links represent an expense, and they may be subjected to enemy jamming or may serve as sources to be tracked by the enemy for the purpose of physical destruction. The design of such data links in a countermeasures environment, is the subject of another paper at this symposium.

This paper addresses a second approach to autonomous target acquisition that is possible against fixed targets, both strategic and tactical, for which prior reconnaissance exists. In this case, a human observer studies the reconnaissance imagery of the target area, finds the target and marks it, and this marked imagery (in some appropriate form) is placed on board the weapon. In addition, the weapon must also carry an imaging sensor, similar or perhaps identical to that required to be on board in the data-link approach, for obtaining "live" imagery during the approach to the target area. If these two images can then be brought into coincidence, the weapon will "know" where it is with respect to the target and will be able to steer itself in. So the technical problem reduces to one of comparing two pieces of terrain imagery that are similar, but certainly not identical, and that are presumably at least partially overlapping. The acquisition function involves answering, with some pre-

assigned level of confidence, the questions "Do the two images in fact overlap, and can they be brought into registry?" The tracking function corresponds to measuring the displacement between the centers of the two images after registration has been consumated, and repeating this measurement with updated sensor images obtained as the weapon approaches the target. These displacements are then supplied to the guidance system to effect terminal guidance and "homing" onto the target. This process is colloquially referred to as "map matching", or sometimes as image correlation--although, as we shall see, correlation in the strict mathematical definition of that word may or may not be required.

If we had weapons with this kind of guidance in the operational inventory today, they would have several important advantages. Primarily, of course, there is the attainment of useful levels of accuracy without exposing manned aircraft to severe defenses; but, by the same token, this system should only be used when the expected aircraft attrition forces us to do so. Next, this system functions without the use of data links, and this has several implications. Not only is the possibility avoided of losing effectiveness due to enemy jamming of the links but, on the positive side, first, the electronic circuitry needed to perform the map-matching probably costs less and certainly weighs much less than the antennas, transmitters, power supply and circuitry of a 2-way data link, second, this system is passive (nonradiating) and so gives the enemy less warning and less opportunity for tracking and attacking either the weapon or its launch platform, and third, it provides the aircrew with a true launch-and-leave capability. Finally, this system, like the data

link systems, utilizes an on-board sensor which also can be jammed or deceived by decoys and camouflage; however, this system, by virtue of its redundant processing of information from a finite area surrounding the target, is actually quite resistant to the effects of localized jamming and to even drastic alterations to the target itself or to discrete portions of the scene.

Despite all these potential advantages, there are no operational systems of this sort in existence today, and for two very good reasons. The most obvious difficulty is the need for prior reconnaissance; and this means not just any old recce picture of the area (which would be a stiff constraint even if that were all there was to it)--it is much worse than that. Ideally, the recce should be taken with the same sensor (same sensitivity and resolution), from the same point in space, at the same time of day, in the same weather, during the same season, etc., as for the strike mission. This is clearly an impossible requirement to meet; but fortunately it is not an absolute requirement since, as pointed out at the end of the previous paragraph, the system is tolerant to some errors. Nevertheless, the point is that all differences between the recce picture and the "live" sensor image will degrade the map-matching performance, so that a good deal of effort must be expended in adapting the raw recce picture to correspond to the predicted strike conditions.

The second difficulty with image correlation systems is more subtle but probably even more important. The problem is that, under the degraded conditions just mentioned, the system may lock onto a secondary peak in the correlation function, producing a "false" match with some

similar but incorrect portion of the target area. Of course, there is no one around to holler "Tilt!" and no way for the system to correct itself, so the result is a gross error in the guidance of the weapon and a wasted mission. This problem, although not encountered frequently, has still been the nemesis of most of the experimental systems in the past that have been brought up to a flight test.

Lest the realization of this kind of guidance appear completely hopeless at this point, I have here a simple demonstration that will illustrate the basic principles. [Note: this consists of some special transparencies which the author will bring to the symposium; they require only a vu-graph or 8-1/2 x 11 transparency projector.] The first pair of transparencies are identical except that one is the negative of the other. As the projectionist superimposes them and slides one over the other, you will observe the marked dip in the total amount of light transmitted when the two images are in register. The second pair of transparencies have identical geometry, being simultaneous photographs of a portion of greater Los Angeles, but one is taken with yellow light and the other with near-infrared radiation. The differences are real, specific and repeatable, not random like TV snow. Upon superimposition of these images you can observe that the null position is less pronounced than in the first case, yet you would undoubtedly still pick the correct match position. This illustrates the point made earlier about robustness or resistance to the effects of small errors--up to a point. But if you try to visualize what the additional effect would be if one of the images were at the wrong scale or had other geometrical distortions,

you can see that the principal null would be still further weakened and that soon some of the secondary dips would be just as deep as the main one; and that is all it takes to get a false lock or gross error.

As has already been hinted, the idea of using map-matching for missile guidance has been around for a long time, and indeed the Air Force has actively sponsored hardware development programs in this field for at least 20 years. Various sensors have been used--optics, radar, radiometry, TV and IR. The early implementations were all analog; today most of the work is digital. However, the principles, and the fundamental strengths and weaknesses, are still the same. Because of the past history of flight-test failures, and the fact that most past analyses have concentrated on the much easier question of the achievable tracking accuracy while largely ignoring the gross-error problem, a small group at Rand decided to direct its effort to more fully understanding the acquisition aspects of image correlation. We think we have made a little progress in this direction, and this is the work that I will describe in the remainder of this paper.[1]

---

[1]This work is reported more fully in two Rand reports, R-2057/1-PR, *Image Correlation, Part I: Simulation and Analysis* by H. H. Bailey, F. W. Blackwell, C. L. Lowery and J. A. Ratkovic, November 1976, and R-2057/2-PR, *Image Correlation, Part II: Theoretical Basis* by H. W. Wessely, November 1976.

II. A PRELIMINARY ANALYTIC APPROACH

STATEMENT OF THE PROBLEM

The mathematical formulation of the correlation problem can be introduced as follows. A pre-strike reconnaissance sensor images a scene on the ground containing a target. This image is broken down into an array of M (square) picture elements (sometimes called pixels), and a value X_I, representing a certain level of a gray scale, is assigned to each element. These data are stored in a computer memory and are henceforth referred to as the "reference map". This reference map, shown schematically in Fig. 1, contains the target located at the center. At some time later, another sensor on board an aircraft or a weapon images a smaller portion of this same scene containing N elements, Y_I, which are similarly digitized to form a "sensor map". (In Fig. 1, both maps are assumed to be square, purely for convenience of notation; thus, $m = \sqrt{M}$ and $n = \sqrt{N}$ represent the number of picture elements in one row of the reference and sensor maps, respectively.) The center of the sensor map, which is related to the boresight of the weapon, will generally be displaced from the center of the reference map by some unknown amount that depends on mid-course navigation and pointing errors, and the sensor map may or may not actually contain the target. The first problem, then, is to find (if it exists) the portion of the reference map that matches the sensor map. Once this is accomplished, the displacement or offset between the centers of the two maps (shown with two components [K,L] in Fig. 1) serves as the correction signal to the guidance/control subsystem.

In attempting to properly locate the sensor map relative to the reference map, the sensor map must be compared with numerous equally sized portions of the reference map. In Fig. 2, the sensor map is shown in two positions--one is the correct or matching position, and the other is one of the Q possible nonmatching positions.[2] The displacement from the correct location of the sensor map to any arbitrary position is defined to be the displacement vector, here indicated by a single index, J. In the absence of certain geometrical errors to be discussed later, all elements of the sensor map are correctly positioned with the corresponding elements of the reference map when the displacement vector is zero. The acquisition phase of the image correlation problem thus reduces to a two-state discrimination problem, i.e., to one of discriminating between the case when the displacement vector equals zero (termed the in-register case) and the case when it does not (termed the out-of-register case).

"METRICS" AND FORMULAS FOR P_c

The actual point-by-point comparison of the sensor map with the reference map is made by computing the value of one of several possible functions of the displacement, J. Algorithms suitable for this computation can be considered, for the moment, as arbitrarily selected functions or "metrics" whose efficacy is to be tested empirically. The justification (or lack of it) for some of the possible choices is dis-

[2] When the separation between positions is measured in units of the cell size, this number Q is, in the case of one-dimensional strip maps, simply M - N; for two-dimensional maps, it is $(m - n + 1)^2 - 1$.

cussed more carefully in Section IV of this paper. The most commonly used algorithm is derived from classical correlation, which is approximated by computing finite sums, sometimes normalized, of the form

$$\phi_1(J) = \frac{1}{N}\sum_{I=1}^{N} X_{I+J}Y_I \quad \text{(Prod)}. \tag{1}$$

This is referred to as the Product algorithm. The next most important metric is the mean absolute difference (MAD) algorithm, defined as

$$\phi_2(J) = \frac{1}{N}\sum_{I=1}^{N} |X_{I+J} - Y_I| \quad \text{(MAD)}. \tag{2}$$

When the maps are in register, $J = 0$; and if no errors are present, $X_I = Y_I$. It can easily be shown that $\phi_1(J)$ has a maximum value and that $\phi_2(J)$ has a minimum value under these conditions. Thus, if it can be assumed that for some test position the two maps really do coincide, then the value of J for which $\phi(J)$ is an extremum essentially defines the best match position between the two maps. However, in the presence of noise and various other errors to be discussed, the extremum only defines the correct match point *on the average*. Because of these effects there is only a certain probability, P_c, (over an infinite ensemble of maps) that the extremum actually defines the correct match point. If $p(\phi|S)$ denotes the conditional probability density of the value of the metric when the maps are matched (S = signal present), and if $p(\phi|B)$ denotes the conditional probability density of the value of the metric when the maps are mismatched (B = background present), then for a maximizing

metric the probability of correct acquisition is given by

$$P_c = \int_{-\infty}^{\infty} p(\phi|S) \left[\int_{-\infty}^{\phi} p(\phi'|B)\, d\phi' \right]^Q d\phi, \qquad (3)$$

where Q, as before, denotes the number of mismatched positions. A simple change in the limits of integration describes the probability of correct acquisition for a minimizing metric. $1 - P_c$, of course, gives the probability of a false lock or gross error--the problem mentioned in Section I.

One straightforward, though certainly only approximate, method of calculating P_c has been proposed by Johnson.[3] His method is approximate because of the large number of quite fundamental and admitedly unrealistic assumptions he has made in order to render the analysis tractable. Briefly, he assumes that, with $Y_I = X_I + N_I$, both X_I and N_I are stationary, ergodic and Gaussian distributed, and that X_I, Y_I and N_I are statistically independent. The ratio σ_x^2/σ_n^2 is referred to as the signal-to-noise ratio, S/N.

With these assumptions, it is indeed straightforward, though somewhat tedious, to calculate the ensemble means and variances of various metrics both for in-register (J = 0) and out-of-register (J ≠ 0) conditions, and then to compute P_c by means of a formula of the form

---

[3] Johnson, M. W., *Analytical Development and Test Results Used in Navigation Systems*, AIAA Paper 72-122, presented at the Tenth Aerospace Sciences Meeting (San Diego), January 1972.

$$P_c = \frac{1}{\sqrt{2\pi}\,\sigma_o} \int_{-\infty}^{+\infty} \exp\left[\frac{-w^2}{2\sigma_o^2}\right] \cdot \left[\frac{1}{2} \pm \frac{1}{2} \operatorname{erf} \frac{(\bar{\phi}_o - \bar{\phi}_j) + w}{\sqrt{2}\,\sigma_j}\right]^Q dw, \quad (4)$$

where $\bar{\phi}_o$ and $\bar{\phi}_j$ are the ensemble mean values of the metric when the maps are in and out of register, respectively; σ_o^2 and σ_j^2 are the ensemble variances when the maps are in and out of register; w is $\phi_o - \bar{\phi}_o$; and Q is the number of out-of-register values of J as defined in footnote 2. The quantities $\bar{\phi}_o$, σ_o^2, and σ_j^2 are simply related[4] to N, σ_x^2, and σ_n^2. Values of P_c have been obtained by numerical integration of Eq. (4) for various values of the parameters N, Q, and S/N within the ranges of $10 \leq N \leq 10^4$, $10 \leq Q \leq 10^6$, and $0.1 \leq S/N \leq 30$, using Rand's IBM 370/158 computer.

---

[4] The relevant ensemble statistics for the quantities $\bar{\phi}_o$, σ_o^2, and σ_j^2 are shown in the table below. Corresponding formulas for other algorithms have been derived but are not needed here.

Algorithm	$\bar{\phi}_o$	σ_o^2	σ_j^2
Prod	σ_x^2	$(\sigma_x^2/N)(2\sigma_x^2 + \sigma_n^2)$	$(\sigma_x^2/N)(\sigma_x^2 + \sigma_n^2)$
MAD	$\sqrt{2/\pi}\,\sigma_n$	$(1 - 2/\pi)(\sigma_x^2/N)$	$(1 - 2/\pi)(\sigma_x^2 + \sigma_n^2)/N$

NUMERICAL RESULTS

Some of the results of these calculations are shown in Fig. 3, where P_c is plotted as a function of the S/N ratio for two specific values of N and Q, for both the Product and MAD algorithms. In order to illustrate more clearly the nature of the dependence on the parameters N and Q, additional data are presented in Figs. 4 through 6 in the form of contours of constant probability (P_c = 0.99, 0.90, 0.70, and 0.50) as a function of N and Q for three different values of S/N, again for both algorithms.

CONCLUSIONS

The following significant conclusions can be drawn from these data.

- The probability of correct match, P_c, increases with an increase in the elemental signal-to-noise ratio, S/N, and with an increase in the size of the data sample, N. (This, of course, is to be expected, since the *total* signal-to-noise ratio represented by the sum of the contribution from each sensor map element is increased by either an increase in the elemental signal-to-noise ratio or an increase in the total number of sensor map elements.)

- The probability of correct match, P_c, decreases with increasing Q, the number of out-of-register positions, but the decrease with increasing Q is relatively weak in comparison to the effect of a change in the number of sensor map elements, N, as shown in Figs. 4 through 6. (This behavior can be explained qualitatively by the fact that an increase in the number of out-of-register positions tends to increase the root-mean-square variation in the value of the metric roughly as the square root of Q, whereas an increase in N increases the value of the

metric linearly when the two maps are matched.)

- At low signal-to-noise ratios ($S/N \leq 1$), the Product is the better algorithm, i.e., it leads to higher values of P_c. (This result is analogous to the well-known finding in statistical communication theory that a correlation receiver is the "matched filter", the best receiver for detecting a signal in noise; however, as explained in Section IV of this paper, the map-matching problem is fundamentally different and the apparent analogy cannot be pressed.)

- At high signal-to-noise ratios ($S/N \leq 3$), the MAD is the better algorithm.[5] (A heuristic explanation for this result is also given in Section IV. More importantly, it is shown next in Section III that, in *real-world systems applications*, the high values of S/N that would justify use of the MAD metric are very seldom realized.)

---

[5] When $1 < S/N > 3$, results are mixed and the choice of algorithm is not critical.

III. EFFECTS OF VARIOUS ERRORS AND REAL SCENES

Realism can be added to the foregoing in two distinct stages. Firstly, the analysis can be extended to include the effects of a number of commonly encountered systematic errors--i.e., differences between the sensor and reference maps other than the simple additive noise treated above. Secondly, simulations can be carried out using data from real scenes which exhibit various amounts of spatial correlation and typical non-Gaussian "structure" in the data per se, quite apart from the problems of system implementation.

DISCUSSION OF ERROR SOURCES

In general, there are at least four classes of systematic errors that can degrade correlation performance:

1. Geometrical distortions.
2. Systematic intensity changes.
3. Quantization errors.
4. Enemy jamming.

1. Any *geometrical distortion* of the sensor map coordinates relative to the reference map coordinates degrades, in ways that are discussed below, the performance of a map-matching system. The four most important types of geometrical distortion are synchronization, rotation, scale factor (magnification), and perspective errors. The detailed analysis of these effects, for digital systems, involves synthesizing

a grid of cells each of which is given a value that is an appropriately weighted average of the values of the distorted cells that partially overlap each of the undistorted cells. These errors are illustrated in Fig. 7, where, for each case, the four cells surrounding the center of the reference map are depicted, together with the corresponding cells of the distorted sensor map.

Synchronization errors occur because there is no way to ensure a common origin between the sensor and the reference map grids. As shown in the figure, this type of error results in all the grid elements of one map being fractionally displaced from those of the other map. This displacement can cause each sensor map grid element to overlap as many as four grid elements of the reference map. The effects of synchronization errors are most significant when the dimensions of a grid element are comparable to the average dimensions of a statistically independent scene element. These errors can have roughly the same effect as reducing any value of S/N to less than unity.

Rotation errors can be caused by heading or attitude reference errors on board the weapon. If the sensor map is centered but rotated relative to the reference map, the map-matching process compares a single sensor cell with a combination of fractions of both matching and nonmatching reference cells. The amount of overlap with nonmatching cells increases as one moves radially outward from the center of the two maps.

Uniform magnification or scale errors are primarily caused by errors in weapon altitude or range to the target, although in some cases they

may be caused by several other effects as well. In the presence of scale factor errors, the sensor elements are dimensioned either somewhat larger or somewhat smaller than the reference map elements. Consequently, elements of the sensor map, when overlaid on the reference scene, will again encompass both matching and nonmatching reference elements, with the amount of nonmatching overlap increasing as one moves radially outward from the center. However, if rotation and scale factor are controlled, as they usually can be in practice, such that $n\theta \leq 1$ and $n(\rho - 1) \leq 1$, where θ is the rotation in radians and ρ is the magnification, then performance is not seriously degraded.

Perspective errors occur when the sensor views the reference area from a different position in space, because of midcourse navigation inaccuracies, for example. Owing to the difference in perspective, a grid pattern of square cells is transformed into an array of trapezoids. Thus, the effect is similar to a linearly varying scale factor error.

When geometrical distortions are present, only a partial match between sensor and reference map elements is possible. When the map centers are slightly displaced, some of the previously nonmatching map elements are brought into coincidence, so that a partial match condition holds for these displacements. The overall effects on the correlation function or comparison metric are thus twofold: the peak value of the metric for the matched condition is reduced, and the breadth of the function is increased. That is to say, the chances of a false lock are increased and the tracking accuracy is decreased.

2. *Systematic intensity errors* include all changes in the amplitude (or intensity) of the sensed scene relative to the reference scene that cannot be attributed to sensor noise. These changes can be aggregated into four general categories: (a) uniform change in overall signal level, (b) shadowing and obscurations, (c) changes in scene reflectivity/emissivity, and (d) reference map construction errors.

The overall signal level of the sensor scene relative to the reference scene can be altered by changes in scene illumination (i.e., day/night or sun/overcast) or by changes in sensor gain settings. Changes in the optical properties of the atmosphere can also change the overall signal level and/or the contrast perceived by the sensor. Shadows due to clouds or changes in sun angle, and obscurations due to intervening hills or foliage, cause blocks of sensor data elements to be totally dissimilar to the corresponding reference data elements. The reflectivity of certain portions of a scene can change drastically as a result of physical changes on the ground, such as snowfall or flooding, or less drastically but significantly as a result of differences in moisture content or seasonal changes in foliage and vegetation, or to a still lesser degree simply because of differences in the direction of the illumination by either active sensors or the sun at different times of day. Finally, the sensor scene can be different from the reference scene owing to actual changes in the reference scene (e.g., new man-made objects and to reference map construction errors. This last category includes all errors made in producing the reference map, but primarily refers to errors made in transforming the original reconnaissance data taken in one

spectral region (e.g., photographic) into a reference map for use with a "live" sensor in a different spectral region (such as infrared or millimeter waves).

These systematic errors generally do not increase the width of the correlation function significantly, but they do certainly reduce the differential between the in- and out-of-register values, thereby increasing the possibility for false locks.

3. In digital correlation systems, the sensor data, which are usually analog in nature originally and may have any one of a continuum of values at each pixel, are quantized into discrete levels and encoded. This process gives rise to what is sometimes called (through analogy with photographic systems) gray-level coding errors, or, *more generally*, *quantization errors*. The effects of quantization become important when only a few gray levels are used and when other errors are present simultaneously. Under those conditions, a signal that is distorted or has had noise added to it may either be coded exactly like the original (in which case, the effect of the noise or distortion has been eliminated) or it may be coded at a different level (in which case the effect of the original error is usually exacerbated). Thus, the effect is something like the addition of noise, but it is a peculiar, non-Gaussian, kind of noise.

4. Finally, *enemy jamming* can cause (a) additional noise, possibly time varying, in all or a portion of the sensor elements, or (b) in severe cases, complete saturation of most or all of the sensor elements. As with the "block" errors described above, the principal effect of jamming is usually to weaken the extremum value of the comparison metric, thus de-

creasing P_c.

More detailed results of these analyses can be found in the cited Rand reports.[1] These include a number of quantitative relationships, derived theoretically and confirmed by simulations (see below), that predict the gross error rate as a function of system parameters and the various types of errors. The most important overall conclusion to be drawn is the following: real systems suffer from unavoidable synchronization errors, from geometrical distortions that can be partially controlled (at a cost) but not eliminated, and from amplitude changes, some of which are partially controllable (such as quantization, system malfunctions, and detector noise) but most of which are not controllable (the real changes in the scene). These systems are therefore invariably operated under conditions of rather low effective signal-to-noise ratios. Thus, despite the simplicity and the apparent advantages of a MAD or similar algorithm as discussed in Section II, a Product algorithm will almost always be superior in real-world applications.

SIMULATIONS USING REAL DATA

It was decided to test this general and admittedly over-simplified theory by carrying out computer simulations of the image-correlation process using data from the real world. A data base of digitized imagery was made available, and the simulations were done as follows. A piece of the original data was selected to serve as a reference map. A section out of this map was chosen and modified (corrupted by the addition of random noise, distorted spatially, or what not) to represent a sensor map. Then the two were correlated in the computer, using

both the Product and MAD algorithms, the process being repeated several times using different noise examples, and the number of occasions when a false lock resulted was actually counted. Thus, empirical values of P_c were "measured", and these were compared with the theoretical predictions.

A very interesting result ensued. The simulations never did worse and usually did somewhat better (i.e., fewer false locks) than the theory predicted. When various of the systematic errors were included, the added effect roughly confirmed the theory, and the general result just stated continued to hold.

The explanation probably lies in the fact that the statistics of real scenes are not Gaussian in nature, as was specifically assumed in the theory. That is to say, the specification of just two parameters--root-mean-square signal amplitude and single correlation length--is far from adequate for characterizing terrain. It is obvious that two of the scenes we had chosen, representing agricultural and suburban regions, contain structure (predominant spatial frequencies) that are, of course, far from Gaussian in character. The strong inference persists, supported by intuition and common visual experience, that correlation algorithms by their very nature are capable of exploiting these structural (non-Gaussian) features and thereby achieve high acquisition capabilities. A proper analysis of map-matching performance in the future can be expected to include not only the effects of signal-to-noise ratio and spatial correlation (the Gaussian properties of a scene that have been considered in the foregoing), but also some yet-to-be-determined higher-

order statistical parameters and/or special ad hoc feature descriptors.

A major conclusion can be drawn at this point: an approximate lower bound on the value of P_c--the probability of correct (and autonomous) target acquisition--can be calculated, so that one can, at least in principle, design systems to meet an acquisition specification.

Quantitative relationships have been presented that show the dependence of P_c on N (the sensor map size), M (the search area or reference map size), S/N (nominally the signal-to-noise ratio but, more importantly, a measure of the fidelity of the reference map vis à vis the real-time sensor map), and various parameters describing systematic intensity and geometrical errors. Thus, one has the tools for carrying out design tradeoffs on sensor resolution and field of view (to increase N), on midcourse navigation (to decrease M), on attitude reference and guidance (to reduce geometrical distortions), on data processing capabilities (to reduce both synchronization and quantization effects), on more recent and more accurate reference data (to increase S/N), and so on, including, finally, a tradeoff of the cost of increasing the P_c requirement itself with the loss of those few weapons that will be wasted if they achieve a false lock.

Most of the above-mentioned relationships for P_c are derived from a simple Gaussian theory that is known to be unrealistic. Fortunately, however, this theory appears to err on the conservative side--most scenes are more distinctive than assumed and results are better than predicted. On the other hand, real systems have additional error sources that have not been analyzed in the experiments conducted in this study. The im-

portant point is that, with a "floor" established for P_c, there should be no major surprises in future flight tests of either experimental or operational hardware in the field of image correlation guidance; improvements in the theory, and additional data from simulation experiments using specific scenes of interest, can only improve the predictions and relax some of the design restrictions. One *can* design to P_c requirements, though at the moment not as effectively as would be desired.

IV. A FRESH APPROACH

Part of the Rand effort was to reexamine the fundamental nature of the map-matching problem and to investigate the degree of theoretical justification for the use of the various metrics. It is not possible to condense the substance of that analysis in the time available here, so the serious student is referred to the second reference given in the first footnote; however, the results can be summarized briefly as follows.

Since the problem is basically one in statistical decision theory, it is shown that the optimum solution is achieved by computing the likelihood ratio for each comparison and then choosing the match point at the place where the likelihood ratio is maximum. Unfortunately, that computation requires a knowledge of the N-dimensional *joint probability* distributions--functions that are unknown and, in a practical sense, unmeasurable. Hence, one must resort to approximations. These usually take the form of maximizing or minimizing one of several functions, herein called "metrics". In much current work, these are chosen almost arbitrarily and therefore must be subjected to essentially experimental validation. By considering two-picture-element scenes, such that the likelihood ratio and several of the commonly used metrics can be expressed in simple algebraic form and discussed in geometrical terms, the essential features of the various metrics are explained and compared with the likelihood ratio. In this way heuristic arguments are developed that support the use of the Product algorithm when S/N is low and the MAD algorithm when S/N is high. The notion, borrowed from signal

detection theory, that the classical correlator (i.e., the Product algorithm) is optimum for this application, is shown to be erroneous; the two problems are fundamentally different.

We conclude, firstly, since there is no processor that is necessarily optimum, that we are free to look for other algorithms. Secondly, since the best way to avoid false locks, is to provide a high signal-to-noise ratio, and since one way of achieving a high S/N ratio would be to enhance the characteristic features in a scene and throw away the rest of the data that contributes mostly noise, we conclude that a deliberate effort to search out the unique or distinctive features in any scene would be highly desirable.

An obvious question follows: what constitutes a "characteristic feature?" Or, stated more carefully: what are the invariant properties of a scene--those least likely to change with time or environmental conditions and therefore most likely (in our context) to lead to high values of the signal-to-noise ratio? I think it is clear, following the discussion of errors in Section III, that the useful information resides more in the spatial relationships than in the intensity or signal amplitude dimension of an image. Absolute intensity levels are the least dependable quantities, and intensity ratios and even the algebraic sign of differences (contrast) or gradients are unreliable. But the locations of most intensity boundaries are fixed--subject only to certain geometrical distortions which are (a) limited in magnitude (largely controlled by the system design), and (b) constant or slowly varying over a given image. It is so easy for the human visual system to identify such boundaries

and other "lines" (double boundaries), and to recognize patterns and shapes formed by them even when they are considerably distorted, that it may be surprising to learn that it is difficult to instruct a machine (i.e., program a computer) to do these things digitally; but such is indeed the case.

There exist a number of gradient and Laplacian operators, and line-following and boundary-seeking algorithms; but these all have problems in selecting appropriate scaling parameters (how big a feature to look for), and problems of implementation when applied either singly or in combination to noisy digitized images. Fortunately, in the applications we are discussing for which prior reconnaissance of fixed targets is required, there is (usually) time for experimentation with the recce imagery and selection, on a specific ad hoc basis, of "filters" or special image-processing algorithms appropriate to each scene. In particular, it is suggested that techniques currently being developed in the field of pattern recognition should be explored and exploited for this purpose. But also, at least for the present, one should not rule out completely the use of people to pick what "look like" good features and their relevant filters. In any case, the on-board processing would then consist of applying only the selected filters, looking only for those features known to be significant in that weapon's assigned target area, followed (presumably) by a simple differencing algorithm for locating the match point. It is anticipated that in this way correlation algorithms more efficient than those commonly used to date will evolve, and that at the same time higher values of P_c (i.e., fewer gross errors) will result.

To sum up, the main thrust of this second portion of our work has been to shift the emphasis away from schemes which compare every pixel in the reference and sensor images to searches for ways to extract the invariant and "distinctive" features in any given scene. A small effort is continuing in the development of a more general (and more satisfying) theoretical analysis of this problem; but in the meantime it is felt that some of the ad hoc approaches that have been suggested will lead to useful military systems in the near future. This is now an active area of research, but no one is ready to make a specific system recommendation at this time.

It should perhaps be emphasized once again that this whole paper is addressed to the acquisition phase of image--correlation guidance. It is assumed that pre-programmed software changes will be able to effect the transition to accurate tracking following successful target acquisition.

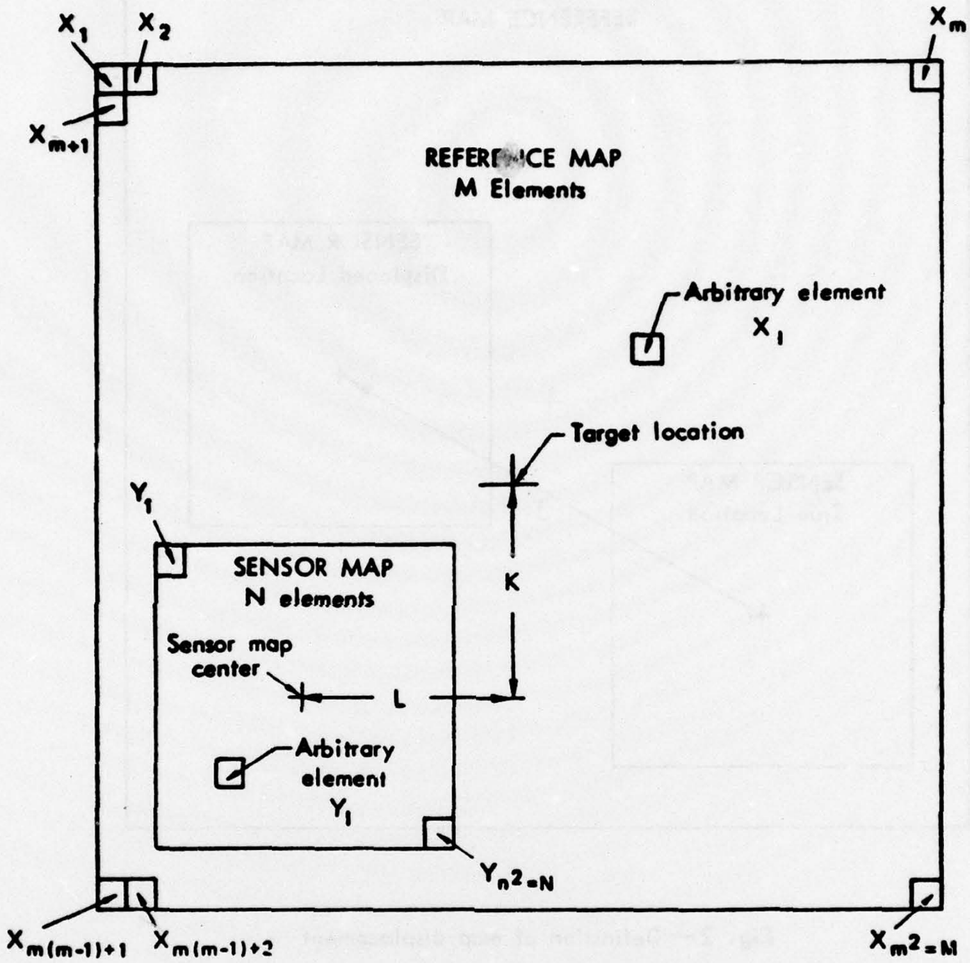

Fig. 1—Map definitions

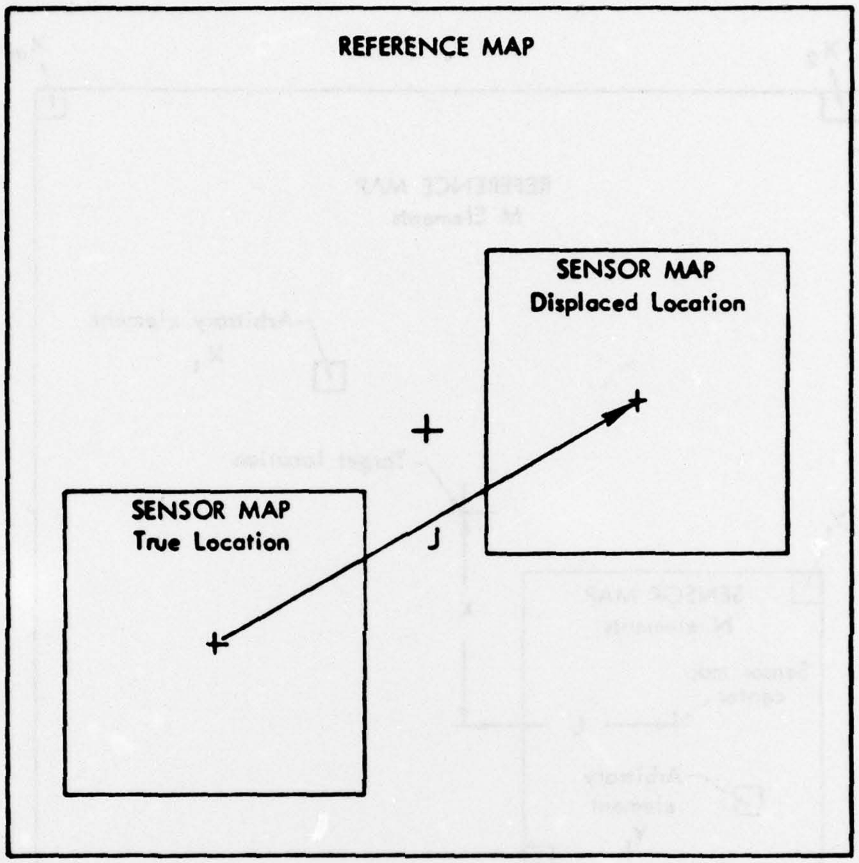

Fig. 2 — Definition of map displacement

N = number of sensor map elements

Q = number of out-of-register positions

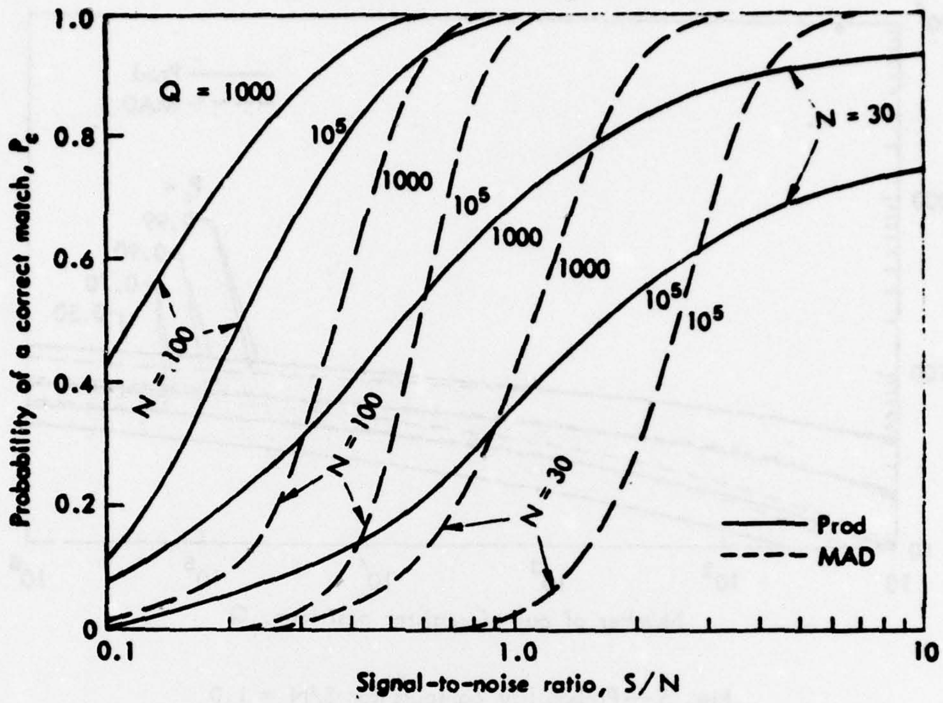

Fig. 3—P_c versus S/N for additive noise only

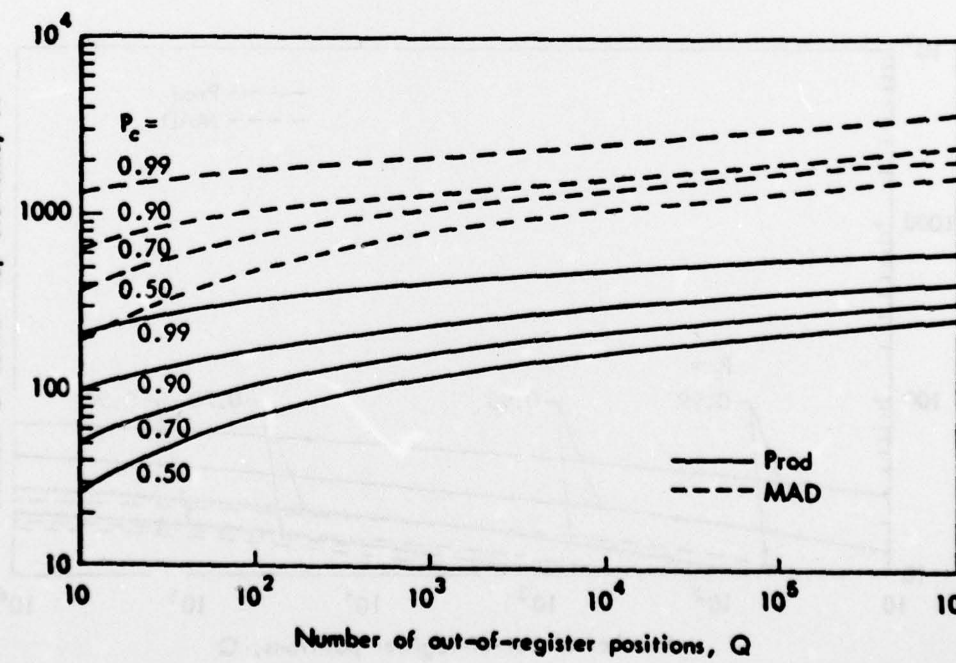

Fig. 4—Probability contours for S/N = 0.1

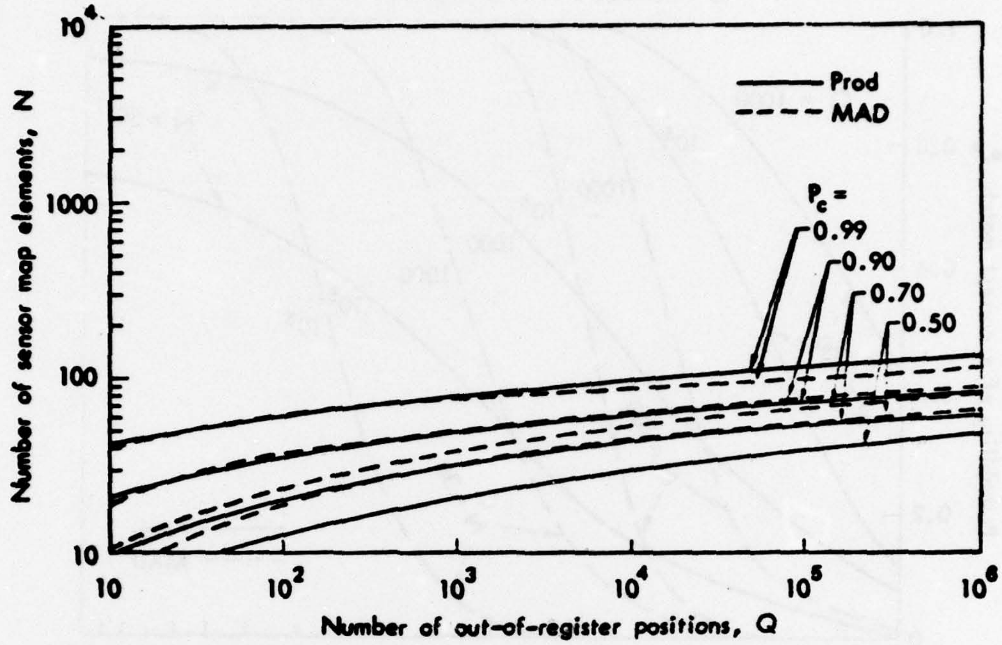

Fig. 5—Probability contours for S/N = 1.0

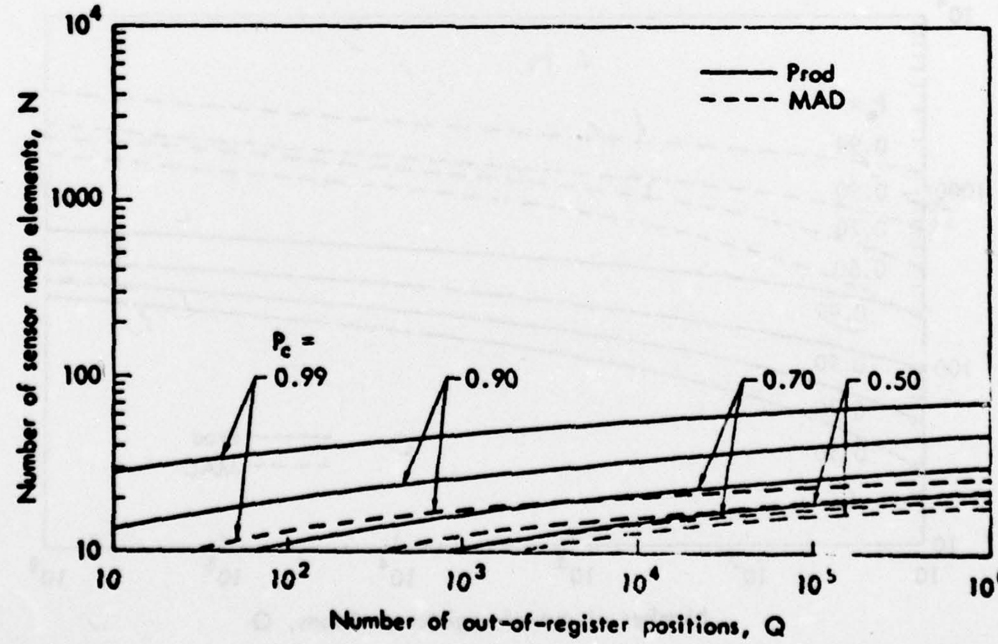

Fig. 6—Probability contours for S/N = 30

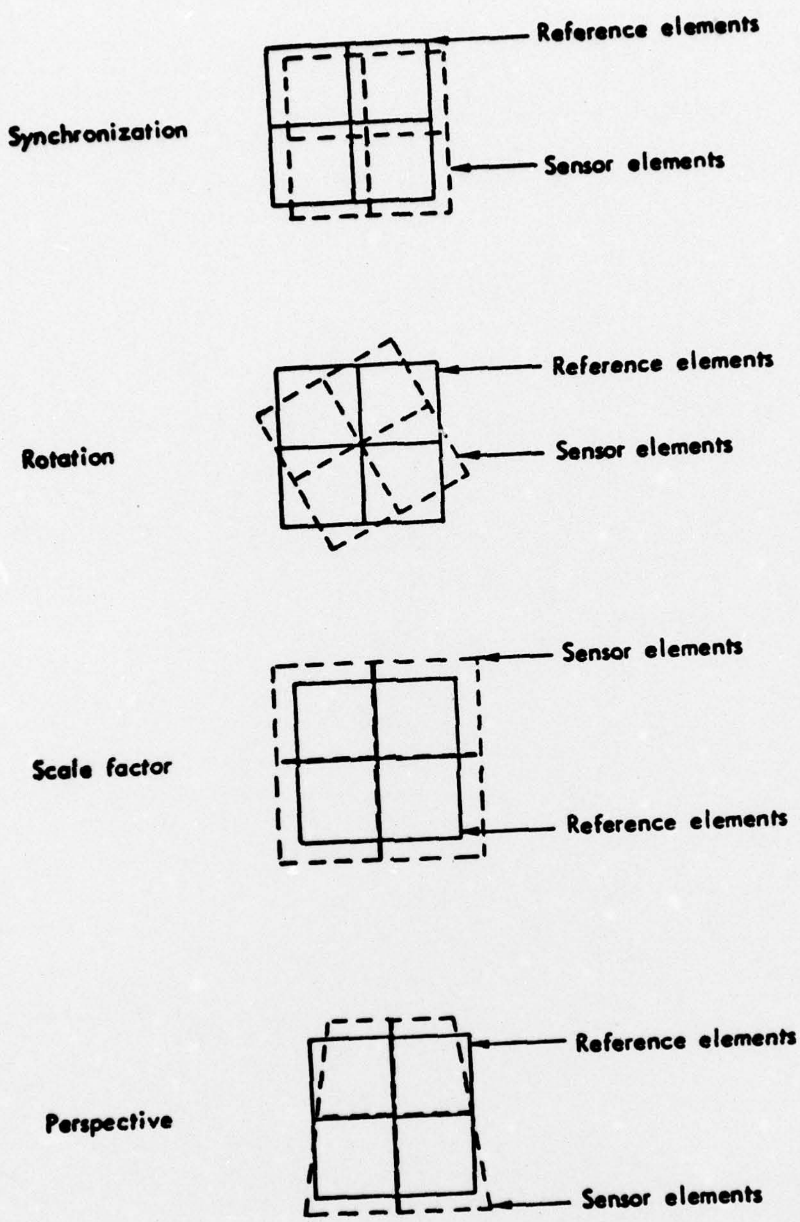

Fig. 7—Geometrical distortion errors